A Brief Introduction to Prox-affine Forms in Convex Optimization

Junzi Zhang, Anqi Fu, Stephen Boyd

EE & ICME Departments
Stanford University
Prox-affine Form of Generic Convex Optimization

We consider the following prox-affine [WWK2015, FZB2019] formulation of a generic convex optimization problem:

minimize \( \sum_{i=1}^{N} f_i(x_i) \)
subject to \( \sum_{i=1}^{N} A_i x_i = b \).

with variable \( x = (x_1, \ldots, x_N) \in \mathbb{R}^{n_1 + \cdots + n_N} \), \( A_i \in \mathbb{R}^{m \times n_i} \), \( b \in \mathbb{R}^{m} \).

- \( f_i : \mathbb{R}^{n_i} \to \mathbb{R} \cup \{+\infty\} \) is closed, convex and proper (CCP).
- Each \( f_i \) can only be accessed through its proximal operator:

\[
\text{prox}_{t f_i}(v_i) = \arg\min_{x_i} \left( f_i(x_i) + \frac{1}{2t} \|x_i - v_i\|_2^2 \right)
\]
Remark on $\text{prox}_{tf}$

- Generalization of projection: take $f(x) = \mathcal{I}_C(x)$, $\text{prox}_{tf} = \Pi_C$.
- $\text{prox}_{tf}$ is a smoothing of $f$.
- $\text{prox}_{tf}(x) = x$ iff $x \in \text{argmin}_x f(x)$.
- For many $f$, $\text{prox}_{tf}$ is closed-form.
- $\text{prox} t \sum_{i=1}^N f_i(x_i)(v_1, \ldots, v_N) = (\text{prox}_{tf_1}(v_1), \ldots, \text{prox}_{tf_N}(v_N))$. 
Why **prox-affine** form?

- **Separability**: suitable for parallel and distributed implementation.
- **Black-box proximal**: suitable for peer-to-peer optimization with privacy requirements.
- **Compact representation**: alternative to **conic** standard form.
  - Cone programs can be represented in prox-affine form by consensus without complication (but NOT vice versa).
  - With log, exp, det involved, prox-affine form is much more compact.
Conic Form as Prox-affine Form

**Conic form:** \((\mathcal{K} \text{ is a nonempty, closed and convex cone})\)

\[
\begin{align*}
\text{minimize} & \quad c^T x + \frac{1}{2} x^T Q x \\
\text{subject to} & \quad Ax = b, \quad x \in \mathcal{K}.
\end{align*}
\]

- Used in most solvers: CPLEX, MOSEK, Gurobi, SCS, ECOS, OSQP.
- Target standard form of most modeling languages: CVX*, YALMIP.

**Prox-affine form of cone programs via consensus:**

\[
\begin{align*}
\text{minimize} & \quad c^T x_1 + \frac{1}{2} x_2^T Q x_2 + \mathcal{I}_\mathcal{K}(x_3) \\
\text{subject to} & \quad Ax_1 = b, \quad x_1 = x_2 = x_3.
\end{align*}
\]

- Consensus already used in most conic solvers, so no complication.
- Can be further parallelized when \(\mathcal{K} = \mathcal{K}_1 \times \cdots \times \mathcal{K}_m\).
Portfolio optimization with transaction costs and risk constraints:

maximize \[ r^T x - \sum_{j=1}^{n} (a_j |x^j| + b_j |x^j|^{3/2}) \]
subject to \( (w + x)^T \Sigma (w + x) \leq \rho, \quad 1^T x = 0. \)

Here \( x = (x^1, \ldots, x^n) \) (and similarly for other variables hereafter).
Example: Portfolio Optimization

**Conic form:** (e.g., what MOSEK, GUROBI, SCS & ECOS accept)

\[
\begin{align*}
\text{minimize} & \quad -r^T x + \sum_{j=1}^{n} (a_j t_1^j + b_j t_2^j) \\
\text{subject to} & \quad \sum^{1/2}(w + x) = g, \|g\|_2 \leq \alpha, \alpha = \sqrt{\rho}, \mathbf{1}^T x = 0, \\
& \quad x^j + t_1^j \geq 0, x^j - t_1^j \leq 0, x - z \leq 0, x + z \geq 0, \\
& \quad (z^j)^2 \leq 2s^j t_2^j, (w^j)^2 \leq 2v^j u^j, z = v, s = w, u = \frac{1}{8} \mathbf{1}, \\
& \quad s \geq 0, t_2 \geq 0, v \geq 0, u \geq 0, j = 1, \ldots, n.
\end{align*}
\]

Variables: \( x, t_1, t_2, g, \alpha, z, v, s, w, u \) – dimension = \( 9n + 1 \).

- complicated transformation and redundancy.

- **more consensus variable copies** are implicitly created to *separate the projections* in the solvers (e.g., SCS) – dimension \( \geq 9n + 1 \).
Example: Portfolio Optimization

**Prox-affine form:** (Epsilon & a2dr)

\[
\text{minimize} \quad -r^T x_1 + \sum_{j=1}^{n} a_j |x_2^j| + \sum_{j=1}^{n} b_j |x_3^j|^{3/2} + I_{\|x_5\|_2 \leq \sqrt{\rho}}(x_5),
\]

subject to \( \sum_{j=1}^{n} w + x_4 = x_5, \ 1^T x_1 = 0, \ x_1 = x_2 = x_3 = x_4 = x_5. \)

Variables: \( x_1, x_2, x_3, x_4, x_5 \) – dimension = 5n.

- Straightforward & compact: more dramatic with log, det, exp.
- Separation into low dimensional problems (easy parallelization): \( \text{prox}_t \sum_{i=1}^{N} f_i(x_i)(v_1, \ldots, v_N) = (\text{prox}_t f_1(v_1), \ldots, \text{prox}_t f_N(v_N)). \)
- Closed-form \( \text{prox}_{tr^T} x_1, \text{prox}_{ta_j} x_2^j, \text{prox}_{tb_j} x_3^j|^{3/2} \) and \( \Pi_{\|x_5\|_2 \leq \sqrt{\rho}} \).
- No additional consensus variable copies are needed: can be directly solved by Epsilon & a2dr.
Related Solvers and Links

Epsilon (2015)
- expression tree compiler for transforming convex optimization problems into prox-affine forms

POGS (2015)
- first-order GPU-compatible solver for graph form convex optimization problems; graph form is similar to prox-affine form

a2dr (2019)
- (Anderson) accelerated Python solver for prox-affine distributed convex optimization
  - https://github.com/cvxgrp/a2dr.
